DOCUMENT RESUME

ED 201 340 IR 009 439

TITLE	Individualized Instruction for Data Access (IIDA). Quarterly Report No. 8.
INSTITUTION	Drexel Univ., Philadelphia, Pa. Graduate School of Library Science.; Franklin Inst. Research Labs., Philadelphia, Pa.
SPONS AGENCY	National Science Foundation, Washington, D.C.
PUB DATE	Mar 80
GRANT	DSI-77-26524
NOTE	72p.; For related documents, see ED 190 080-081.
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	Comparative Testing; *Computer Assisted Instruction; Computer Programs; Databases; *Individualized Instruction; *Information Retrieval; Online Systems; Questionnaires; *Search Strategies; *Training Methods; Use Studies
IDENTIFIERS	*Individualized Instruction For Data Access

ABSTRACT

A brief summary of the progress and status of the Individualized Instruction for Data Access (IIDA) project is followed by a report dealing with two studies in the environment for which the system was intended. These field experiments were conducted with research engineers and chemists at two different Exxon Research and Engineering Company sites. In one of these--the baseline study--the performance and reactions of IIDA trained and assisted researchers were compared with those who had searches done for them by professional intermediaries. In the second--the training method study--the performance and reaction of IIDA trained and assisted searchers were compared with those of searchers trained by conventional methods and then assisted in searching by IIDA. Thus the focus in the first study was upon the adequacy of IIDA as a system for retrieving information, while the second was upon IIDA as a way of providing online search training. Findings are discussed as they pertain to the goal of providing a method of allowing direct user access to bibliographic searching, and recommendations are made for future research. Copies of both the pre- and post-searching questionnaires are appended. (RAA)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED201340

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

INDIVIDUALIZED INSTRUCTION FOR DATA ACCESS (IIDA)

Quarterly Report No. 8
March, 1980

Drexel University, School of Library
and Information Science
Franklin Institute Research Laboratories

NSF Grant No. DSI 77-26524

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Paul Meadow

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

IR009439

TABLE OF CONTENTS

1. Overview
2. Baseline Study
    2.1 Procedure
        2.1.1 Design
        2.1.2 Selection of participants
        2.1.3 Scheduling of participants
        2.1.4 Conduct of the study
        2.1.5 Data coding
    2.2 Results
        2.2.1 Demographic information
        2.2.2 Diagnostics
        2.2.3 Post-search questionnaire
        2.2.4 Responses of IIDA users to open-ended questions
    2.3 Discussion
3. Training Method Study
    3.1 Procedure
        3.1.1 Design
        3.1.2 Selection of participants
        3.1.3 Scheduling of participants
        3.1.4 Conduct of the study
        3.1.5 Data coding
    3.2 Results
        3.2.1 Demographic information
        3.2.2 Diagnostics
        3.2.3 Post-search questionnaire
        3.2.4 Responses of IIDA trained users to open-ended questions
        3.2.5 Responses of conventionally trained users to open-ended questions
    3.3 Discussion
4. General Discussion
5. References
6. APPENDIX A: Questionnaires

1. Overview

This project was a renewal of earlier work on Individualized Instruction for Data Access (IIDA). Begun in July, 1976, with initial funding for one year, the project was resumed in April, 1978, and is effectively completed with this report and the final report which follows.

The project staff were divided into two groups. The computer group has been concerned with the design, implementation and testing of the requisite computer programs. These various programs have been described in detail in an earlier report (1). From the user's point of view, there are four major subsections of the total system--three instructional exercises and the on-line search assistance mode.

By the end of February, 1979, the first and second exercises had undergone system debugging. Preliminary formative evaluation was conducted in early March by various kinds of computer literate users actively looking for flaws in the system. With the second exercise providing the nucleus which was expanded into the assistance mode, both the third exercise and the assistance mode were ready for use in evaluation testing in the summer of 1979.

The behavioral group of the project staff has been concerned with both formative and summative evaluation of IIDA. In formative evaluation, the concern has been with monitoring system development and providing feedback and information for refinement and development of the system. In summative evaluation, the concern has been with the assessment of the impact and effectiveness of IIDA, and with the extent to which the objectives of the project have been met. The preceding quarterly report (2) discussed the series of formative evaluations conducted with IIDA, and also indicated some of the kinds of changes made as a result of the evaluations. In addition, that report also dealt with one set of evaluations which were both summative and formative in nature in that they were concerned both with the question of the impact of the system on the user and with modification of the system for future use. Both the results of these studies and some of the changes made in the system as a result of the studies are described in that report.

The substance of this report deals with two studies conducted in the environment for which the system was intended. During the fall of 1979, and the winter of 1979-1980, two field experiments were conducted with

research engineers and research chemists at two different Exxon Research and Engineering Company sites. In one of these studies--the baseline study--the performance and reactions of IIDA trained and assisted searchers were compared with those of end users who had had searches done for them by the professional intermediaries at the Information Center on site. In the second study--the training method study--the performance and reactions of IIDA trained and assisted searchers were compared with those of searchers who were given more conventional classroom training and then assisted in searching by IIDA. Thus, in one case the focus was upon the adequacy of IIDA as a system for retrieving information, and in the other the focus was upon the IIDA training materials as a way of providing on-line search training.

Before beginning a discussion of the details of the two studies, there are some acknowledgements which should be made. First of all, thanks are owed to P. L. Dedert, E. M. Holbrook, M. K. Landsberg, P. A. Lorenz, M. MacFarquhar, R. Parker, K. E. Shenton, S. J. Swetnick, and K. R. Walton of Exxon Research and Engineering Company, who assisted in setting up and conducting the experiments. Thanks are also extended to the participant users of the Exxon information retrieval services at the sites where the studies were conducted. Without their cooperation and assistance the studies could not have been completed. In addition, especial thanks are owed to Barbara Lawrence of the Exxon Corporation. Without her initiative, the studies reported here would never have been done at all. Finally, the authors would like to thank Robert Rich of Princeton University who served the IIDA project as a consultant on evaluation. The support for the computer time used in the studies described below was provided by the Exxon Research and Engineering Company, which paid the computer costs, as well as by DIALOG and COMPENDEX, who allowed the use of their facilities for the experiments.

2. Baseline Study

The baseline study was conducted at the Exxon Research and Engineering Company facility at Florham Park, New Jersey, in the fall of 1979. The study design called for collecting data on the searches performed by twenty-five searchers who were trained and assisted by IIDA. In addition, twenty-five comparable searches were to be done by the information retrieval staff on site. The intent here was to use the intermediated searches to provide a profile of diagnostic usage for comparison with the diagnostic usage of IIDA users. It was anticipated that significant differences between the two groups of searches would reveal deficiences in the IIDA training and assistance programs which would then be subsequently corrected by appropriate modifications of the system.

2.1 Procedure

Planning for the study at Florham Park began well before the first IIDA staff member visited the facility. Preliminary arrangements were made in August of 1979; the actual on-site participant use of IIDA began the first week in October, and concluded during the third week in December.

2.1.1 Design. The design of this study consisted of a simple treatment and control group experimental design with subjects being randomly assigned to each group. Those participants who were assigned to the IIDA group were trained to search using IIDA programs and then did two searches on subjects of their own choosing. The other participants, in the intermediated group, submitted the next search topic arising from their regular work to the information center intermediaries. These searches were then conducted by the intermediaries through IIDA, but with all of the IIDA diagnostic messages being suppressed. Thus, records of the number of times the diagnostics were triggered were kept by IIDA, but the intermediaries did not get messages from IIDA. All of the searches were done in COMPENDEX.

2.1.2 Selection of subjects. The staff of the library/information center at Florham Park sent out some 2,200 letters to the engineering, management, and professional level personnel on the site, briefly describing the opportunity to be involved in the IIDA bibliographic search training and inviting them to participate. Nearly 150 responses were received either by in-person, written, or telephone contacts. The names were randomly ordered so that date of response, level of position within the company, and

department and section would not be considerations for selection. From this randomized list, twenty-five participants were randomly assigned to the IIDA group, and twenty-five to the intermediated group.

Those volunteers who were not selected to participate initially were notified that they were being put on a list of alternates and would be called if possible. Very early, it became clear that the vagaries of scheduling and system availability would make it difficult for all of the participants to complete the requisite portions of the study within the time period allotted for the conduct of the study; consequently, two alternates were added to each group. By the end of the study, intermediated searches had been done for twenty-six participants, and twenty-four of those assigned to the IIDA training and assistance group had completed the full course of training and searching. In all cases, the participants were unable to complete the necessary parts of the study as a result of scheduling problems or conflicts with travel to other Exxon sites.

2.1.3 Scheduling the subjects. Considerable difficulty was encountered in scheduling the participants in the Florham Park study resulting from the fact that the use of IIDA was limited during the first month of the study. This limitation was imposed by unexpected hardware restrictions on the computer which housed the IIDA software, as a result of which, we were restricted to the hours between noon and 5:00 p.m. Since most of the participants left the site by 4:30, the half-hour after that was utilized by the intermediaries for completion of the searches submitted by the intermediated group. During the second month of this study, the hardware problem was eased, and IIDA could be used for the entire day. Consequently, the restrictions on scheduling for the second month were those of staff and participant availability.

During the first four to five weeks, during which the intermediated searches were done, the IIDA training group participants were the only ones scheduled for training. All participants were originally scheduled for two sessions. The first appointment was scheduled for one hour, during which it was expected that the participants would complete both exercises one and two. The second appointment was scheduled for two hours, during which the subjects were expected to complete exercise three, and their two searches. This particular scheduling pattern was selected as a result of

expectations that trying to do everything in one sitting was too much, and that the participants would find it difficult to fit more than two appointments into their normal work schedules. Every attempt was made to schedule the first and second appointments within a period of no more than one week. In many cases, however, a third appointment was necessary as a result of technical problems or as a result of either late arrival or early departure on the part of the participants.

2.1.4 Conduct of the study. Each of the fifty participants was sent a pre-search questionnaire before beginning the experiment, with identical questionnaires being used for each group of participants. The participants were asked to return the completed questionnaires to the IIDA staff either by mail or when they came in for their first appointment. No user was allowed to begin IIDA training without having returned a completed questionnaire. When the participant arrived for the first appointment, the project staff took the questionnaire and then filled out a card on the participant listing name, office phone and address, password, and the schedule of appointments which had been set up for the participant. For the intermediated participants, a card was completed at the time the intermediary completed the participant's search, which was done after the completion of the pre-search questionnaire. The card for these participants listed the information mentioned above, as well as the password used in completing the intermediated search for the participant. The cards for both groups were color coded to indicate whether the participant was a member of the IIDA or the intermediated group, and were also used to record the portions of the training completed at each session, the amount of time it took to complete each exercise and each search, and to note any additional appointments or any kind of help sought by the participant from the project staff rather than IIDA. Any peculiarities of the participant's IIDA experiences were noted on the card.

The IIDA staff agreed in advance upon certain "rules" for monitoring the subjects' activities. Very little attention had been paid in the development of IIDA to creating a foolproof, user-oriented procedure for logging on to the system initially, or for recovering from DIALOG or telecommunication crashes. Consequently, an IIDA staff member was physically present on-site for all IIDA training and searching sessions so as to deal with these problems. This solution to the problem of access to the system,

however, created other problems in that with another person in the room, or nearby, some participants tended to want to use that person as a help library. The procedures established for dealing with user questions directed to IIDA staff basically fit the following hierarchy. If a participant asked for help from a staff member, the staff member made every effort to direct the participant to IIDA's help programs to answer the question. If the participnat persisted and did not want to use the help program, or did not know how, the staff member would further encourage the use of help or briefly remind the user of how to get into the help program. When all such efforts failed, the question was answered as briefly as possible and as directly as possible, and then a note was made of the incident on the participant's card. There were very few questions which the staff had to answer directly. Most of the questions directed initially to the staff involved how to get started with the first practice search (exercise two), or how to decide which optional material to take a look at in exercise three. These questions to the staff were easily deflected back to IIDA. The participants also asked frequently about the TYPE formats, but were easily led to the help program for a solution to the problem.

The sequence of events for the IIDA trained participants was as follows: (a) complete the pre-search questionnaire; (b) complete exercises one, two, and three, and the first search; (c) complete the intermediate questionnaire; (d) complete the second search; and, (e) complete the post-search questionnaire. The transcript of each participant's exercise one was given to him upon completion of the exercise. Transcripts of exercises two and three, and of the individual searches, were Xeroxed, and the originals returned to the participants.

For the intermediated participants, the sequence was as follows: (a) complete the pre-search questionnaire; (b) submit a search topic to the information center staff; and, (c) complete a "post-search" questionnaire about the results of the intermediated search. Transcripts of the intermediated searches were returned immediately to the participants along with the questionnaire about the search results. The post-search questionnaire filled out by the intermediated participants consisted of the first page of questions from the intermediate questionnaire filled out by the IIDA trained participants. These questions were duplicated in the IIDA group post-search questionnaire. Samples of the questionnaires used in this study appear in Appendix A.

Since all of the participants were recruited with an offer of bibliographic search training, those who were assigned to the intermediated group were given the opportunity for IIDA training after their intermediated searches, and the questionnaires about them, had been completed. Scheduling of training sessions for these users began early in the second month of the study. For the eighteen who actually completed the IIDA training, all records were kept, and they were also asked to complete the post-search questionnaire used with those participants who had been assigned to the IIDA training group initially. These data are not reported on herein.

2.1.5 Data coding. Using a pre-established coding procedure, each questionnaire was coded, on the questionnaire, as it was completed by the participant. As each participant's folder of questionnaires and transcripts was completed, the folders were returned to Drexel, where the coding was checked by another staff member while the data was transferred to data sheets to be used for keypunching. These sheets were then checked by yet another staff member as the data cards were punched. Finally, the data cards were checked against the original coding on the questionnaires. Diagnostic and search history data recorded by the IIDA record keeping programs were periodically listed out at MIT and sent to Drexel, where the errors, etc., diagnosed by the IIDA diagnostics were compared with the Xeroxed transcripts of the searches. This provided a check on the possibility of system or programming errors resulting in erroneous data recording. In the few cases where the data was lost from the machine records as a result of system failure, the data was taken directly from the search transcripts. When all of this data had been transferred to coding sheets, the data cards were punched, proofed, and corrected if necessary. These cards were then checked against the original machine records and/or search transcripts, and corrected if necessary.

2.2 Results

Basically four kinds of data will be discussed below. These are: (a) demographic information collected on the pre-search questionnaires; (b) the frequencies with which the various categories of machine diagnostics were activated; (c) the various measures collected on the post-search questionnaire; and, (d) a description of the remarks made by the IIDA users on the post-search questionnaires.

2.2.1 Demographic information. With the exception of two participants, one in each group, all of those who took part in the Florham Park

study were engineers of one type or another. None of the participants had ever done on-line searching before, but almost all had had on-line searching done for them at some time prior to the beginning of the study, some as often as eight or nine times. The overall average number of times the participants had had on-line searching done was almost exactly three, and the two groups did not differ in this respect. All but five participants in each of the two groups knew at least one computer language. The two groups did not differ in the number of languages known, and the overall average was approximately 1.5. In addition, the two groups did not differ in their ratings of their level of computer skills, with the overall average rating being 2.8. The two groups also did not differ in their pre-search attitudes towards using the computer for work related projects. This was assessed using a series of six bi-polar adjective scales scored from one to seven, with one being least favorable. The average total score for each group of participants was approximately thirty.

2.2.2 Diagnostics. The diagnostic procedures utilized by IIDA in providing search assistance represent both ways of deciding when a user needs assistance, and an index of how often a particular difficulty occurs. The diagnostics are a set of rules, clustered into categories, which are used by the computer to decide which messages to send to the user, and when to send them. When a diagnostic routine detects a problem, IIDA intrudes on the user with a statement of what the problem appears to be, and with an indication of how to get advice or reference information. Thus the frequency with which a particular diagnostic is triggered represents the number of times that a particular problem arose in the searching behavior of the user.

One type of diagnostic is concerned only with the validity of a DIALOG command, and is thus entirely context free. The guiding operational rule is that a command is valid if and only if it would be accepted by DIALOG. The second type of diagnostic--the local or individual command usage diagnostics--look at the most recent command in the context of the accumulated history of previous commands. These rules deal with problems that range from fatal errors (such as the use of an undefined set number in a COMBINE command) to mere inelegance of usage (such as repeating a previous command). The specific sub-categories of this kind of diagnostic utilized by IIDA in this study include those that deal with: (a) repeated commands,

(b) creation of null sets, (c) unused sets, (d) uninformative print formats, (e) excessive printing, and, (f) excessive time to enter a command.

The global, or command string diagnostics--the third major category--deal with a set of commands as an entity. They are, in effect, concerned more with style than with mechanical exactitude. The first sub-category of this type of diagnostic consists of rules which deal with string and cycle length. Basically, a string of commands consists of one or more commands of the same type--e.g., SELECT, SELECT--and a cycle basically consists of some number of strings of commands in which the user creates a set of references, takes a look at some of the members of that set, and then cycles back to refine that set. Generally, with novice users, an unusually long string of commands is an indication of a user problem. Two other types of diagnostics in this global category deal with either thrashing or dwelling. Thrashing involves changing the "direction" of the search rapidly or often, without any apparent progress toward a goal. Dwelling involves remaining with a search concept when it may well be time to give up and take another approach. The last sub-category of diagnostics in the global category deals with relevance. These diagnostics are used to direct the searcher's attention to the fact that a particular set seems not to be fruitful, or that a previously examined set yielded higher relevance scores.

Before reporting the results of the comparisons between the searches conducted by the intermediaries and those conducted by the IIDA trained searchers, there is an important point which must be discussed. As the result of an IIDA program system difficulty, some of the search records for the intermediated searches were not preserved by IIDA and printed out for later use. This problem was unrelated to the actual conduct of the searches by the intermediaries, but it was not discovered until many of the search transcripts had already been returned to the participants. Since several of the participants had already made use of, and disposed of, the search transcript, several sets of records were completely lost. Of the twenty-six search topics actually submitted to the intermediaries, IIDA diagnostic data were available on only sixteen of the searches. Since only sixteen sets of search records were available, the comparisons between the two groups on the diagnostic measures had to be conducted on fewer searches than originally intended. In addition, the possibility exists that the

sixteen remaining sets of diagnostic information were not representative of what the group of records would have looked like had the entire set been available. This raises a concern about whether or not it is reasonable to compare the diagnostic data of the intermediated and IIDA trained groups even though ten sets of diagnostic data are missing from the intermediated group.

While this question can not be answered with certainty, there are two factors which make it plausible to assume that the pattern of results reported below does not differ appreciably from the pattern that would have been found if the missing sets of data had been included. First, one of the participant identification numbers indicated the order in which the searches were done for the participants, and a Mann-Whitney U test was conducted comparing the two groups of intermediated searchers, and treating the identification numbers as data. Second, the post-search questionnaire measures of those participants whose diagnostic data were missing were compared with the measures for those whose diagnostic data were not missing. For all but two of these latter comparisons, the Mann-Whitney U test was used. For the other two--the two dichotomous variables--Fisher's exact test was used. Since it is not desirable in this case to have a difference between the two groups be undetected, the conservative alpha level of .25 was adopted. None of the tests mentioned above indicated a significant difference between the two groups of intermediated participants. In fact, the smallest probability attained by any of the test results was greater than .32, so that it seems very unlikely that differences exist between the two groups of intermediated participants on the measures tested.

The results mean two things. First, the problem which resulted in the loss of diagnostic data on some of the intermediated searches was unrelated to the order in which the searches were done. Second, the reactions of the participants to the search results produced by the intermediaries did not differ for the two groups who had intermediated searches done. Thus, the effect of the loss appears to be equivalent to the effect expected if the ten sets of diagnostic data had been dropped entirely at random. While this does not guarantee equivalence on the diagnostic data between the searches where the records are and are not available, it does imply that the comparisons reported on below, between the remaining intermediated searches and those done by the IIDA trained participants, probably would

have led to the same conclusions even if the diagnostic data for ten of the intermediated searches had not been lost as the result of a programming error.

In looking at the comparisons between the searches done by the intermediaries and those done by the IIDA trained searchers, the pattern is one in which there appear to be no differences between the two sets of searches. The mean frequencies for each group of searches for each of the three major types of categories of diagnostics--syntax, individual command usage or "local" diagnostics, and command string usage or "global" diagnostics--are illustrated in Figure 1. For each of these classes of diagnostics, the differences between the two searches are not statistically significant (Mann-Whitney U test, $p > .05$). This finding also holds true for each of the sub-categories of diagnostics. These data are summarized in Table 1 for the intermediated searches, and in Table 2 for the searches done by the IIDA trained searchers. It should be noted that there is one sub-category of diagnostic information which is reported in Table 2 but is not included in the results illustrated in Figure 1, and which was not included in the comparisons conducted between the two groups. Since the intermediaries did their searches with IIDA suppressed, they were not asked by the system to make relevance judgments on the references typed out. Consequently, this diagnostic category was not included in the data of the IIDA users when comparisons were conducted between the IIDA assisted and the intermediated searches.

2.2.3 Post-search questionnaires. In the case of the intermediated group, the post-search data collected is based upon the first page of questions in the Intermediate Questionnaire in Appendix A. This one page was the questionnaire, along with a cover page, which these participants received when sent the results of the search which had been performed for them. The corresponding data for the IIDA trained group comes from the identical questions which appeared in the Post-Search Questionnaire (see Appendix A), which was completed by these users after they had finished their second search. This set of questions includes five which ask the individual to rate his degree of satisfaction with various aspects of the search--amount of effort, amount of time, kind of assistance, items retrieved, and total search experience. Two other questions asked the participant to rate the usefulness of the results of the search, and the

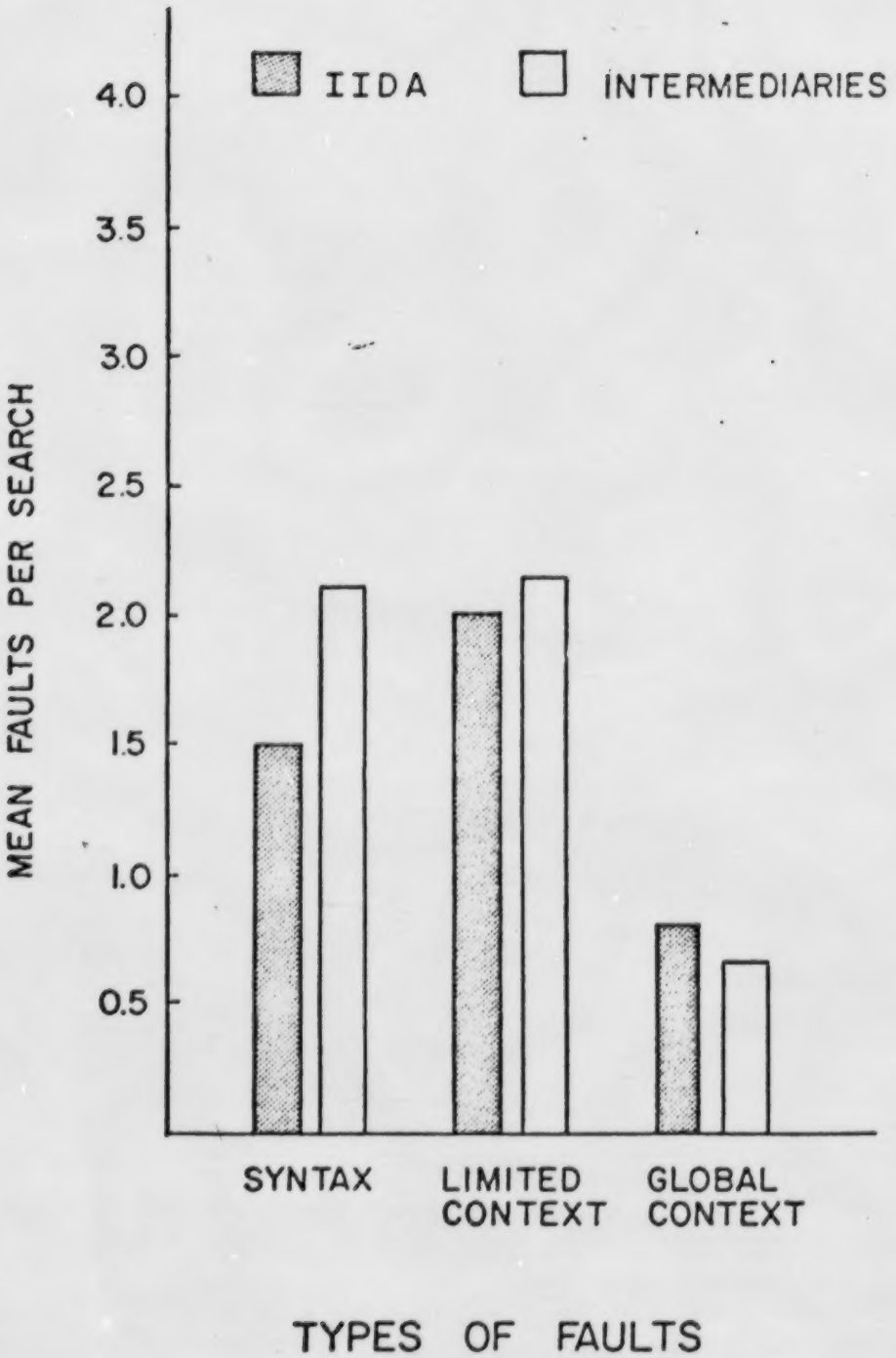

Figure 1. Average frequencies for three major classes of faults at Florham Park.

Diagnostic Category	Total	Avg.	Std. Dev.	Median	Range
1. Syntax	34	2.12	2.33	1.50	0 - 8
2. Command Repetition	1	0.06	0.25	0.03	0 - 1
3. Uninformative Formats	0	--	--	--	--
4. Null Sets Created	14	0.88	1.67	0.17	0 - 5
5. Unused Sets	3	0.19	0.54	0.07	0 - 2
6. Time	1	0.06	0.25	0.03	0 - 1
7. Viewing Requests Excessive	16	1.00	1.10	0.75	0 - 3
8. String/Cycle Length	9	0.56	0.89	0.23	0 - 2
9. Thrashing	0	--	--	--	--
10. Dwelling	2	0.12	0.50	0.03	0 - 2
11. Relevance[1]	--	--	--	--	--
12. Local (2 + 3 + 4 + 5 + 6 + 7)	35	2.19	2.37	1.83	0 - 8
13. Global (8 + 9 + 10 + 11)	11	0.69	1.20	0.23	0 - 4
Total (1 + 12 +13)	80	5.00	4.16	5.50	0 - 11

[1]These diagnostics not operating for these searches.

Table 1. Diagnostic data summary for Florham Park intermediated searches.

Diagnostic Category	Total	Avg.	Std. Dev.	Median	Range
1. Syntax	35	1.46	1.38	1.71	0 - 5
2. Command Repetition	1	0.04	0.20	0.02	0 - 1
3. Uninformative Formats	2	0.08	0.28	0.04	0 - 1
4. Null Sets Created	37	1.54	2.02	0.70	0 - 7
5. Unused Sets	0	--	--	--	--
6. Time	0	--	--	--	--
7. Viewing Requests Excessive	9	0.38	0.58	0.25	0 - 2
8. String/Cycle Length	20	0.83	1.68	0.42	0 - 8
9. Thrashing	0	--	--	--	--
10. Dwelling	0	--	--	--	--
11. Relevance	48	2.00	2.28	1.20	0 - 9
12. Local (2 + 3 + 4 + 5 + 6 + 7)	49	2.04	2.10	1.36	0 - 7
13. Global (8 + 9 + 10 + 11)	68	2.83	3.48	1.50	0 - 14
Total (1 + 12 + 13)	152	6.33	4.67	4.50	0 - 18

Table 2. Diagnostic data summary for Florham Park IIDA trained searcher searches.

usefulness of the search as a whole. In addition, the participants were asked to decide what percentage of the references retrieved were: (a) Very Useful, (b) Useful, (c) Useless. For purposes of data analysis, these first two categories were combined into a single category--percentage useful. The remaining two questions asked the participant to indicate whether or not he had done any previous research on the search topic, and whether or not he would recommend the system to friends.

The data from two of these questions--satisfaction with amount of effort and satisfaction with amount of time involved--were not useable since several of the participants in the intermediated group chose not to answer them. Comparisons were made, however, between the intermediated group and the IIDA trained group on the remaining data. These comparisons were conducted on the sum of the remaining three satisfaction rating scales, on the sum of the usefulness rating scales, on the percentage of useful references retrieved, on the number of users reporting previous research on the search topic, and on the number of users who would recommend the search assistance system to friends. All comparisons between the two groups were conducted either with the Mann-Whitney U test, Fisher's exact test or the Chi square test.

The two groups did not differ from each other in terms of their average total rated satisfaction and their rated usefulness of the search. The intermediated participants gave the searches done for them average satisfaction and usefulness ratings of 9.4 and 5.5 respectively, while the corresponding figures for the IIDA trained and assisted searchers were 9.0 and 5.4, respectively. In addition, the two groups did not differ significantly in terms of their judgments of the percentage of useful items retrieved. The intermediated search participants indicated, on the average, that 49.3% of the references retrieved were useful, while the IIDA searchers indicated that 52.5% of the references were useful.

A total of 85% of the intermediated search participants were willing to recommend that form of search assistance to their friends, while a total of 88% of the IIDA trained and assisted searchers were willing to recommend IIDA to their friends. This difference is not statistically significant. In addition, with twelve of the intermediated and fifteen of the IIDA group having done some previous research on the topic of the search being rated, there was no difference between the two groups in terms of this measure of prior familiarity with the topic.

The intercorrelations of several measures from the post-search questionnaire, the diagnostic categories, and the pre-search questionnaire are reported in Table 3 for the intermediated search group, and in Table 4 for the IIDA trained and assisted group. Not surprisingly, in both groups the individual user's self-rating of computer skills is correlated with the year in which he received the Bachelor's degree. What is not clear, however, is why the IIDA trained searchers' pre-search attitudes toward computer use in a job-related context should correlate significantly with rated computer skill and degree year, when no such correlations appear to exist in the intermediated group. However, even though the two correlations for the IIDA group are significant, they do not differ significantly from the correlations in the intermediated group. Thus, they may be a function of chance rather than of some systematic effect.

The post-search questionnaire measures of satisfaction with the search system, usefulness of the search, and willingness to recommend the system to friends are all significantly intercorrelated for both groups. In addition, the correlations do not differ significantly from each other from one group to the other. However, the satisfaction and usefulness measures both correlate significantly with the judgment of the percentage of useful references retrieved in the IIDA group, but not in the intermediated group. Testing the differences between the correlations from one group to the other indicated a non-significant difference between the two groups for the correlation between usefulness and percentage useful, but a clearly significant difference between the two groups for the correlation between satisfaction and percentage useful ($z = 2.89$, $p < .01$). This suggests the possibility that the number of useful items retrieved is more important in influencing satisfaction with a search assistance system when the user is actually spending some time in doing the search.

There are two other significant correlations in Tables 3 and 4. For the intermediated searches done by the Exxon professionals, there is a correlation between local faults--individual command usage--and global faults--command string usage. For the IIDA group there is a correlation between the global faults and the individual's willingness to recommend the search assistance system to others. However, neither of these correlations differs significantly from the corresponding correlation in the other table, and consequently they may well be a function of chance.

	1	2	3	4	5	6	7	8	9	10
1. Computer Use	1.00									
2. Computer Skill	.20	1.00								
3. Degree Year	.02	.50**	1.00							
4. Satisfaction	-.32	.04	.11	1.00						
5. Usefulness	-.23	.06	.07	.60**	1.00					
6. Recommend[1]	-.32	.08	-.16	.44*	.40*	1.00				
7. Global Faults[2]	.40	-.03	-.35	-.20	-.27	-.29	1.00			
8. Local Faults[2]	.08	.24	-.16	-.16	-.17	.02	.49*	1.00		
9. Syntax Errors[2]	.26	.41	.17	.19	-.23	.01	.11	.13	1.00	
10. % Useful Items	.26	-.10	.15	-.09	.18	.19	.09	.22	-.26	1.00

*$p \leq .05$

**$p \leq .01$

[1]Correlations involving this variable are point-biserial correlations.

[2]Correlations involving this variable are based on an N of 16 rather than an N of 26.

Table 3. Correlations for intermediated search group.

	1	2	3	4	5	6	7	8	9	10
1. Computer Use	1.00									
2. Computer Skill	.40*	1.00								
3. Degree Year	.40*	.62**	1.00							
4. Satisfaction	-.12	-.16	-.18	1.00						
5. Usefulness	.15	.18	.14	.72**	1.00					
6. Recommend[1]	.14	.12	.14	.47*	.39	1.00				
7. Global Faults	-.32	-.16	-.27	-.05	-.05	.65**	1.00			
8. Local Faults	.29	.22	.26	-.13	.09	.13	-.07	1.00		
9. Syntax Errors	-.05	-.07	-.38	.07	.03	.13	.22	.13	1.00	
10. % Useful Items	-.31	-.25	-.14	.66**	.51*	.31	.08	-.24	-.27	1.00

*$p \leq .05$

**$p \leq .01$

[1]Correlations involving this variable are point-biserial correlations.

Table 4. Correlations for IIDA trained search group.

2.2.4 Responses of IIDA users to open-ended questions. On the post-search questionnaire filled out by the IIDA trained and assisted users, there were questions which asked for written comments about various aspects of the search assistance system. While this information was not collected from the intermediated group, and consequently no direct comparisons can be made, it is being presented here, nonetheless, for its overall value in interpreting the findings discussed above.

a) Positive impressions. When asked to give their positive impressions of IIDA, many users reported finding the overall system easy to use and to understand, noting in particular the fast response time. One user referred to IIDA's method of instruction as "simplified coaching," and another felt that the three initial exercises provided sufficient preparation to use the system satisfactorily. Some of the participants also appreciated the ability to select the degree of detail in the explanations in exercise three. The most frequent response was related to the availability of the help library, and IIDA's directions to specific areas of the help library following error messages. For example, one user reported that this feature saved him from making several additional errors. Also referring to the help features of IIDA, another user found the system to be, "...fairly forgiving, even mistakes are tolerated." Other positive impressions dealt with the command summary at the end of exercise one, the optional summary at the end of a search, and about IIDA's tracking system for the relevance of the search. All but four of the IIDA users had had a search done for them by the information center professional searchers at some time prior to the beginning of the study. Some of these users favored their experience with IIDA. As one user put it, "I can retrieve information quickly doing my own search because I know better what I am looking for." Others commented about the convenience of doing their own searching, and about the elimination of the middleman."

Several user comments were directed toward the positive aspects of online searching in general. Some users mentioned the simplicity of searching by creating sets in steps, which was helpful in narrowing one's search topic, and one user felt that EXPAND was especially helpful in this way. Another participant noted that the EXPAND command was very useful in providing a source of ideas for alternative searches. Other features the users liked included the ability to retrieve descriptive abstracts and the ability to screen citations before asking for more details (various TYPE formats).

b) Negative impressions. When asked to give their negative impressions of IIDA, several users reported having trouble remembering all of the instructions, and having difficulty in looking back over the printed exercises. Most of them added they would have preferred a small "primer" type manual with a brief summary of major steps, commands, and type formats. Some even felt this primer would have been just as good as IIDA on-line. Concerning IIDA's instructions, one user thought there should be more assistance ("coaching"), while another found IIDA "too verbose."

A number of the responses dealt with the TYPE format and command. One user reported that TYPE format six gave too much information. Others suggested that there should be a means of terminating the printout on a citation if one finds the first part printed useless. A few users were dissatisfied with the limit on the number of citations which could be requested. Two of the participants expressed aggravation with the relevance ratings they were asked to make following each typed record. As one said, "It drives you crazy."

Many of the comments dealt with various aspects of DIALOG or the data base. For example, several users appear to have had trouble selecting the correct descriptors in the data base to retrieve their specific requests. They suggested that IIDA ought to provide an "automatic list of related terms," separate from the EXPAND command, as they appear in the file. Most of these users felt that additional assistance in this area would have helped them in narrowing down their searches. Other comments included: "Some articles didn't seem to fit the category for which I was searching," "Articles appeared which I thought I had already excluded from the final set." A few users complained that the combination of sets did not seem complete. Others disliked the importance placed on the order of the descriptors (e.g., "concrete blocks" retrieves different items than "blocks concrete), and the need to have separate sets for singular and plural forms of a descriptor. Only one user felt that the available data base was inadequate.

Some of the participants were more concerned with the mechanics of the system, and found the print difficult to read, or the information printed too slowly for "first pass" searching. Several other negative comments concerned the mechanical or telecommunication problems which developed, such as, "...frequent inability to make connections, "...interruptions in program execution," and noise on-line causing the system not to accept all commands.

c) IIDA did the wrong thing. When asked to list situations in which IIDA did the wrong thing, the majority of responses concerned telecommunication problems such as: delays logging on, general noise on-line, disconnections from DIALOG in mid-search, delays in receiving commands, and some communications breakdowns resulting from improper responses. In one case, IIDA was clearly at fault, in that the system ignored the user's request for a summary report at the end of the search: "It just logged off."

d) Search without IIDA. If the participants felt they had learned enough about searching that they could search on their own, they were asked to say when they first decided this. Of the sixteen who responded, three decided after exercise two; three reported deciding after exercise three; one, during the first search; five, after the first search; two, during the second search; and two, after the second search.

When asked to explain their reasons for feeling they could search without IIDA assistance, most users felt that the only thing they would have difficulty with was the log on procedure, and getting back on line in cases of telecommunication difficulties. Approximately one-third of these users found the search methods and application of commands easy to learn, and one user added that by the end of exercise three, one accumulates enough material and experience to search without assistance. A few participants reported that they could search without IIDA, but really were not familiar enough with the vocabulary terms to do an efficient search. One user suggested that his search efficiency might improve with some exposure to the professional searchers. Although another user felt he could search without assistance, he admitted that he would prefer a supplementary manual to, "reinforce what I learned from IIDA." Six of the users who responded either misunderstood the question (i.e., to search without assistance), or did not consider the assistance mode part of IIDA. They all reported they could search without assistance, but their explanations included relying on IIDA's help. Some of these responses included: "IIDA will help if needed," "...instructions are always available," "...if problems came up you could always ask for help," etc.

Those users who did not feel that they could search without IIDA were quite uniform in their reasons. They explained that they needed more practice with commands and/or advanced techniques to be able to search with any efficiency.

2.3 Discussion

In this study two randomly assigned groups of users provided feedback on the results of a bibliographic search. For one group of users, the search was done by a professional intermediary. The other group did their own searching with the assistance of the IIDA assistance programs, after having been trained in searching by the IIDA training exercises. Comparison of the two groups on several pre-search measures revealed no significant differences between the groups, indicating that the randomization had had the desired effect. The two groups also did not differ from each other on several post-search measures of their reactions to the searches. For example, in terms of judgments of the percentage of useful items retrieved in the searches, the difference was not statistically significant. This leads to the conclusion that the IIDA trained and assisted users were able to do searches with results equivalent to those produced by trained and experienced intermediaries. It was also found that there were no significant differences between the two sets of searches in terms of the number of IIDA detected errors or faults in the searching behavior of the professionals and in the searching behavior of the IIDA assisted users. Although comparable information was not available for both groups of users, the responses of the IIDA trained and assisted users to open-ended evaluation questions seemed to clearly indicate a positive response to the system by that user group.

In light of the intent of this study, i.e., to provide a set of diagnostic benchmark criteria against which to assess the performance of the IIDA trained and assisted searchers, it was surprising to discover a lack of significant differences between the two groups of searches. IIDA was designed to produce acceptable results for its users, but it was never expected that the step-by-step performance would match that of professionals. A significant difference on one or more of the diagnostics would have pointed toward a deficiency of one kind or another in the IIDA training, or in the usefulness of the diagnostic messages during IIDA assisted searching. This would have led to revision of one or both aspects of the system. The finding of no significant differences also poses a potential problem in that this might suggest either that the IIDA training and diagnostics worked well, or that they are totally irrelevant.

In looking at the lack of differences between the two sets of searches, there are some circumstances which suggest that the professionals may have

been somewhat handicapped relative to their habitual mode of searching. For example, the Exxon professional searchers are used to working with 1200 baud terminals. With these terminals, more on-line searching is acceptable than with the 300 baud terminals used for IIDA searching. Hence, the professional staff was accustomed to doing a great deal of on-line printing, often exceeding the limit allowed for IIDA users working with the slower terminal. Also, IIDA's design was fixed just before Lockheed announced the new SUPER SELECT command (3). Hence this command was not included in the IIDA training materials, and its syntax was not recognized by IIDA's parser.

Although the intermediaries knew about these limitations, an inadvertent transferring of habits from their usual search context into the IIDA search context could have introduced some faults or errors which might not otherwise have occurred. For example, if a searcher used SUPER SELECT for set combination, then there would be increased likelihood of the appearance of a large number of SELECT commands, without any COMBINE command. This could then trigger an excessive string length diagnostic. Consequently, although it is not possible to assign numerical values, there is reason to believe that the error-fault rate of the professional searchers might have been higher than normal. However, it is certainly the case that the IIDA searchers performed respectably well, and, in terms of the final outcome, the users who did their own searching using IIDA achieved results equivalent in utility to those who worked through intermediaries.

In fact, one reason for arguing that the training and diagnostic assistance routines worked as intended lies in the fact that the end user evaluation of the utility of the information retrieved in the searches did not differ significantly between the two groups. This means that a group of individuals who had never before done on-line searching were able, with IIDA training and assistance, to do searches which produced results containing as much useful information as was contained in the results of searches done by professional searchers. They also did not differ significantly from those participants who had had intermediated searches done for them in terms of either their degree of satisfaction with the search system used, or of their willingness to recommend the search system to their friends. Under these circumstances, the notion that the diagnostics are irrelevant as measures of searcher performance seems implausible, as does the notion that the training and assistance programs were ineffective.

3. Training Method Study

The second major test of IIDA took place at the Exxon Research and Engineering Company facility in Linden, New Jersey, during the winter months of 1979-1980. This study involved a comparison of the searches done by IIDA trained and assisted searchers with those done by a group of searchers who were also assisted by IIDA, but who had been trained in a half-day classroom training session. The intent of this study was to test out the instructional procedures and materials involved in IIDA training against a more frequently used and more conventional training methods. Presumably a human instructor is more responsive to the problems which a student may encounter, and presumably the student has a greater range of questions which can be asked of a human instructor, especially since IIDA was not designed to be a question answering system. Consequently, it was anticipated that significant differences between the two groups of searchers would reveal deficiencies in the IIDA training which would then be subsequently corrected by appropriate modifications of the system.

3.1 Procedure

Preliminary planning for this study was begun at the same time as the planning for the Florham Park study. In fact, one of the contingency plans, which remained a plan only, was to use the Linden site as a back-up in the event of an unforeseen difficulty at the Florham Park site. This, however, was unnecessary, so during the third week of December and through the first week of February, users at the Linden site took part in an entirely different study than the one conducted at Florham Park.

3.1.1 Design. The design of this study consisted of a simple two-treatment experimental design with subjects being assigned randomly to each group. Those participants who were assigned to the IIDA group were trained to search using the IIDA programs, and then did two searches on subjects of their own choosing. The other subjects, in the conventional group, were trained to search in a classroom lecture-discussion format, with some time allowed for hands-on search experience, by one of the Linden information center professionals. They were then scheduled to do two searches of their own choosing through the IIDA assistance mode, with IIDA fully operational. All searches were done in COMPENDEX.

3.1.2 Selection of subjects. The Exxon information center staff sent out a letter to the entire research staff at the Linden research

facilities, explaining the opportunity to be involved in bibliographic search training, and inviting them to participate. The letter included a tear-off section for the volunteer to send back with name, section, office and phone number. It also listed the weeks during which the study was to run, one or more of which the volunteer was to circle, indicating availability. Nearly 150 responses were received. They were divided into groups by the information center staff according to the weeks the volunteers were available for the study, and within these grouped according to the section in which the volunteer worked. From these groups, a stratified random sample of fifty names was chosen, so as to obtain an even mix of participants from different sections, and from those available during each of the six to eight weeks during which the study was expected to run. From this list of fifty, twenty-five subjects were randomly assigned to the IIDA group, and twenty-five to the conventional group.

During the course of the study, it became necessary to make a few changes in these groups as a result of schedule conflicts, or of participants dropping out of the study. Two of the participants, originally selected for the conventional group, were switched to the IIDA group because they were unable to attend the scheduled training session. In these cases, two participants, previously selected for the IIDA group, who were available for the conventional training session were switched to replace them. Throughout the course of the study, five participants (three IIDA and two conventional) dropped out as a result of lack of time or of travel requirements. These participants were replaced from the remaining pool of volunteers by making a random selection from the slips for a particular week when a replacement was needed. Finally, using the same procedure as for replacements, one extra participant was added to each group about midway through the study simply to ensure that at least twenty-five participants in each group would finish in the time allotted for the study. When the study was completed, there was, in fact, one participant in each group who had not finished the full course of training and searching.

3.1.3 Scheduling of subjects. During the Linden study, the only restrictions on scheduling participants were the hours of staff and participant availability. Generally, IIDA usage was almost continuous from 9:30 a.m. to 4:30 p.m., Monday through Friday, from December 17, 1979 to February 6, 1980, with five days off over the weeks of the Christmas and New Year's holidays.

All of the participants were scheduled for a total of two-and-one-half hours, including one-and-one-half hour of training and one hour of searching. However, while the IIDA group participants spent the full time on the IIDA system, the conventionally trained participants spent only the one hour of searching on the IIDA system. In general, the participants were scheduled for their various sessions a week or two in advance of actually beginning. All exercises and searches on IIDA were assumed to require approximately one-half hour each. The IIDA participants were all initially scheduled for three sessions, two for an hour each, and the third for a half-hour. The conventionally trained participants were scheduled for one training session lasting an hour-and-a-half, and then for two half-hour sessions on IIDA. The sessions were broken up this way so as to provide the participants some time between each to allow them to assimilate the information learned from each session. Each participant was originally scheduled so as to be able to complete the entire course of training and searching within three days to a week. However, some participants took longer than a week between starting and finishing. These delays resulted from unanticipated schedule conflicts on the part of the participants, or from system problems--e.g., crashes, very long delays in getting logged on, etc.--which required rescheduling.

3.1.4 Conduct of the study. Each participant filled out a pre-search questionnaire before beginning the experiment. These questionnaires were mailed out to all of the participants along with the letter informing them that they had been selected to participate in the study. The participants either returned the questionnaires by mail, or brought them to their first session. None of the participants were allowed to begin the training without having completed a questionnaire. As with the Florham Park study, when a participant arrived for his first session with IIDA, a card was filled out on him with his name, phone and office number, IIDA password, and his appointment dates and times. The training session date for each of the conventionally trained participants was also noted on the card.

The conventionally trained participants took part in one of four training sessions developed and conducted by the Exxon staff. These sessions were scheduled on the 17th of December, 1979, and on the 2nd, the 14th, and the 21st of January, 1980. The classroom portion of each ran from 1:30 to 3:00 in the afternoon, and the number of participants taking part in each session

ranged from five to eight. Although each session was conducted by a different professional, the same instructional materials and handouts were used in each. In addition, each session gave the participants some time on-line to practice the techniques they had learned. All of these participants were also given, either in the classroom session or in the brief introduction to IIDA on the system, a short explanation of the IIDA system, particularly the information on cancelling characters or lines with IIDA. The course handouts used in the conventional training sessions consisted of some forty pages covering the various basics of on-line searching, including system commands, data base structure, and search strategy. Many of the pages presented examples of the effects of various commands, type formats, etc. In addition to this material, these users also had access, during training, to the DIALOG information on COMPENDEX, i.e., sample records, retrieval methods summary, and formats available summary. The conventional training sessions were structured to be roughly comparable in content and coverage to the IIDA training sessions (exercises one through three). A description of IIDA training materials can be found in an earlier report (4).

The conventionally trained participants then performed two searches of their own choosing using IIDA. These searches were performed in the IIDA laboratory during two half-hour sessions. The amount of time actually spent by the participant for each search was recorded on his card, along with any comments about the system problems encountered, or questions which the participant may have asked of the IIDA staff member. After the first search, the participant filled out an intermediate questionnaire, and after the second search, a post-questionnaire. All of the questionnaires used in this study appear in Appendix A.

The IIDA trained participants also came to three sessions, lasting a total of two-and-one-half hours, ideally, but their sessions were structured somewhat differently. In the first hour-long session, the participant completed exercises one and two--the first tutorial exercise and a practice search. In the second session, also lasting an hour, the participant worked through exercise three--another tutorial exercise--and the first search of his own choice. At the end of this session, the participant completed the intermediate questionnaire. For the last session, of half-hour length, the participant again did a search of his own choice, and then filled out the post-search questionnaire. The time spent on each

exercise and search were recorded on the participant's card, along with any comments on system problems or questions asked by the participant. As in the Florham Park study described above, the IIDA staff made every effort to deflect user questions to IIDA rather than answer them directly.

3.1.5 Data coding. The data coding and preparation procedures used in this study were essentially duplicates of those described above for the Florham Park study.

3.2 Results

As with the previous study, basically four kinds of data will be discussed below. These are: (a) demographic information collected on the pre-search questionnaires; (b) the frequencies with which the various categories of machine diagnostics were activated; (c) the various measures collected on the post-search questionnaire; and, (d) a description of the remarks made by the IIDA trained and by the conventionally trained users on the post-search questionnaire.

3.2.1 Demographic information. Although one participant in the IIDA trained group did not have a B.A. degree, all of the other users had undergraduate degrees in fields with a scientific or technical orientation. Nineteen of the participants in each group had degrees in either Chemistry or Chemical Engineering. The remaining users had undergraduate majors in such fields as Physics, Biochemistry, and Physical Chemistry. Two participants in the IIDA group, and one in the conventionally trained group, had done some on-line searching for themselves before the beginning of the study. While five of the IIDA group, and six of the conventionally trained group, had never had on-line searching done for them, most of the users had had searching done for them at some time before the beginning of the study. The IIDA group averaged 4.2 searches for those who had had searching done. The average for those in the conventionally trained group who had had searching done was 3.8 searches. The two groups did not differ on this or any of the measures discussed so far. They also did not differ in their pre-search attitudes toward using the computer for work-related projects, with each group having an average total score of approximately thirty.

There was a difference between the two groups in terms of the number of computer languages known. The average for the IIDA group was 1.4, while the average for the conventionally trained group was .8 (Mann-Whitney $U = 208.5$, $p < .05$). The difference, however, does not appear to be a particularly strong one, since looking at only those in each group who knew one

or more languages indicates that these sub-groups did not differ, with an average of 1.9 for the IIDA group, and of 1.5 for the conventionally trained group. Further, while six of the twenty-five participants in the IIDA group, and twelve of the twenty-five in the conventionally trained group, knew no languages, the difference between the two groups is not significant. Turning to the individual's self-rating of computer skills, the two groups did not differ in their ratings of their own computer skills, with the IIDA group having an average of 2.4, and the conventionally trained group, an average of 2.0. Consequently, there seems to be little reason to assume that the IIDA group had a clear advantage in computer skills over the conventionally trained group.

3.2.2 Diagnostics. In looking at the comparisons between the searches done by the IIDA trained and by the conventionally trained users, again the pattern, as in the preceding study, is one in which there appears to be no differences between the two sets of searches. The mean frequencies for each group of searches for each of the three major categories of diagnostics are illustrated in Figure 2. It should be noted that the data in this study do include the sub-category of relevance diagnostics in the "global" diagnostic category, since both groups did their search with IIDA fully operational rather than suppressed, as was the case for the intermediated searches done in the study described earlier. For each of the three categories of diagnostics in Figure 2, the differences between the two sets of searches are not statistically significant (Mann-Whitney U test, $p > .05$). Furthermore, individual comparison of each of the sub-categories of diagnostics indicated no significant differences between the two groups. These data are summarized in Table 5 for the IIDA trained participants, and in Table 6 for the conventionally trained participants.

3.2.3 Post-search questionnaire. For both groups of users, the post-search data collected is based primarily on the questions in the Post-Search Questionnaire in Appendix A. Participants in each group completed this questionnaire after having conducted their second of two self-selected searches. The questionnaire includes several sets of questions. One set of five asked the individual to rate his degree of satisfaction with various aspects of the search--amount of effort, amount of time, kind of assistance, items retrieved, and total search experience. Two other questions asked the participant to rate the usefulness of the results of the search, and

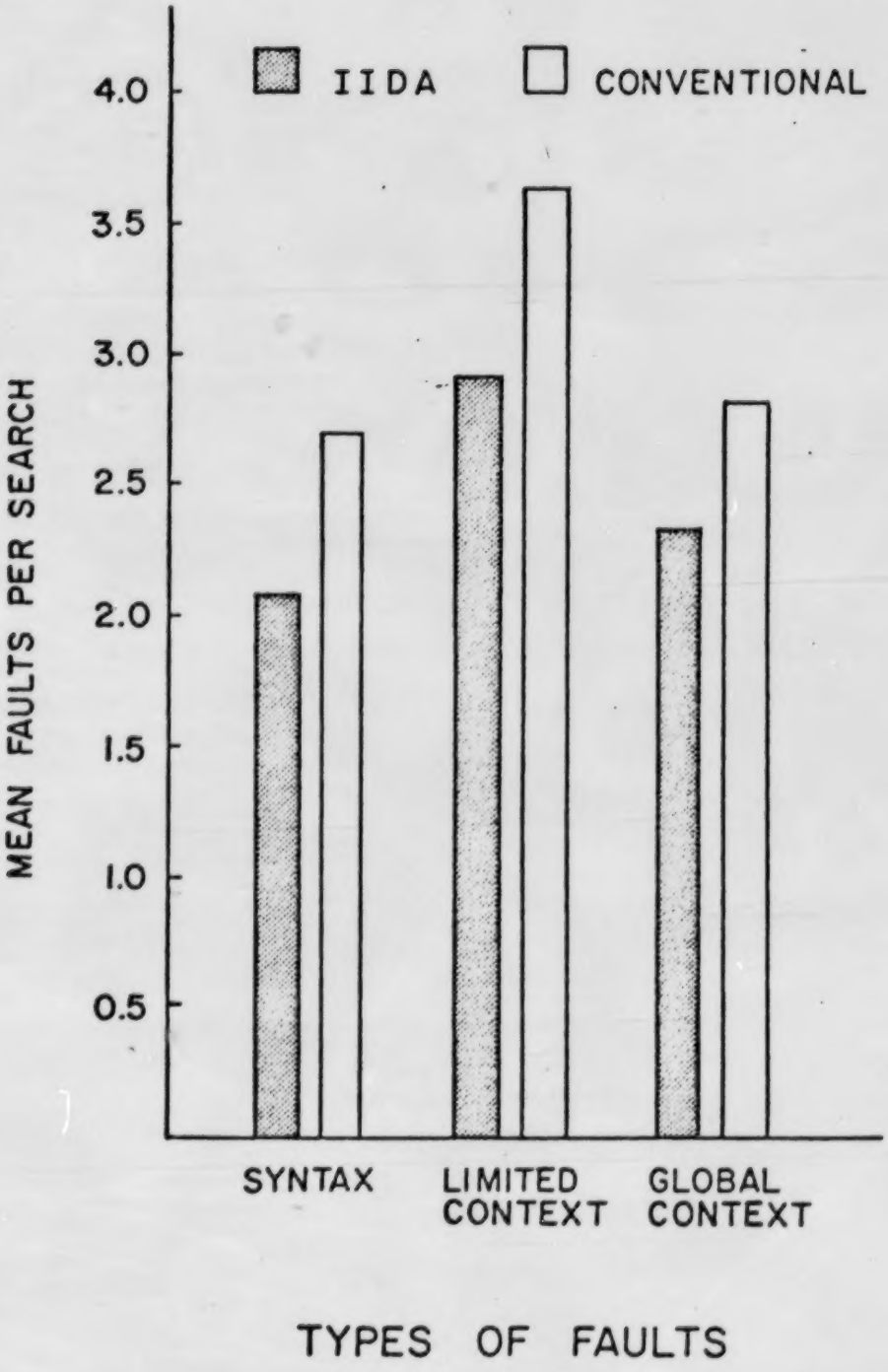

Figure 2. Average frequencies for three major classes of faults at Linden.

Diagnostic Categories	Total	Avg.	Std. Dev.	Median	Range
1. Syntax	53	2.12	1.98	1.43	0 - 6
2. Command Repetition	4	0.16	0.37	0.10	0 - 1
3. Uninformative Formats	4	0.16	0.37	0.10	0 - 1
4. Null Sets Created	40	1.60	2.31	0.39	0 - 8
5. Unused Sets	1	0.04	0.20	0.02	0 - 1
6. Time	13	0.52	1.94	0.04	0 - 9
7. Viewing Requests Excessive	11	0.44	0.65	0.28	0 - 2
8. String/Cycle Length	14	0.56	1.29	0.19	0 - 6
9. Thrashing	0	--	--	--	--
10. Dwelling	0	--	--	--	--
11. Relevance	43	1.72	1.67	1.57	--
12. Local (2 + 3 + 4 + 5 + 6 + 7)	73	2.92	2.98	1.44	0 - 8
13. Global (8 + 9 + 10 + 11)	57	2.28	2.23	1.86	0 - 8
Total (1 + 12 + 13)	183	7.32	5.21	7.00	0 - 20

Table 5. Diagnostic data summary for Linden IIDA trained searcher searches.

Diagnostic Category	Total	Avg.	Std. Dev.	Median	Range
1. Syntax	68	2.72	2.03	2.43	0 - 7
2. Command Repetition	3	0.12	0.33	0.07	0 - 1
3. Uninformative Format	2	0.08	0.28	0.04	0 - 1
4. Null Sets Created	61	2.44	3.44	1.00	0 - 14
5. Unused Sets	2	0.08	0.28	0.04	0 - 1
6. Time	6	0.24	1.20	0.02	0 - 6
7. Viewing Requests Excessive	18	0.72	0.74	0.65	0 - 2
8. String/Cycle Length	32	1.28	1.93	0.39	0 - 6
9. Thrashing	0	--	--	--	--
10. Dwelling	0	--	--	--	--
11. Relevance	37	1.48	1.61	1.09	0 - 7
12. Local (2 + 3 + 4 + 5 + 6 + 7)	92	3.68	3.79	2.62	0 - 15
13. Global (8 + 9 + 10 + 11)	69	2.76	2.76	1.67	0 - 11
Total (1 + 12 + 13)	229	9.16	5.79	9.12	0 - 20

Table 6. Diagnostic data summary for Linden conventionally trained searcher searches.

the usefulness of the search as a whole. In addition, each user was asked to decide what percentage of the references retrieved were: (a) Very Useful, (b) Useful, (c) Useless. Two other questions which also appeared on the first page of the questionnaire asked the user to indicate whether or not he had done any previous research on the search topic, and whether or not he would recommend the search system to friends.

On the last page of the questionnaire appeared a set of questions which asked the participants to evaluate the search system by agreeing or disagreeing with a series of five statements. These items were scored on a scale from one to five, with five indicating a favorable evaluation of the system on that item. Finally, the user was asked to rate the degree of understanding of the information retrieval system gained, and to indicate whether or not he felt that he could search on his own.

Comparisons were made between the IIDA trained and the conventionally trained group on these data. These comparisons were conducted on the sum of the satisfaction rating scales, the sum of the usefulness scales, the sum of the percentages of Very Useful and Useful items retrieved, the sum of the agree-disagree evaluation scales, the rating of degree of understanding gained, the number of users who indicated they had done previous research, the number who indicated they would recommend the system to their friends, and the number who felt able to search on their own. In addition, the two groups of users were also compared in terms of the estimated time required to complete the second search. All comparisons between the two groups were conducted either with the Mann-Whitney U test, Fisher's exact test, or the Chi square test.

The two groups did not differ from each other in terms of their average total rated satisfaction, their rated usefulness of the search, and their gain in understanding of the system. The IIDA trained participants had average satisfaction, usefulness, and understanding ratings of 14.3, 5.1, and 3.4, respectively, while the corresponding figures for the conventionally trained searchers were 15.0, 5.4, and 3.5, repectively. In addition, the two groups did not differ significantly in terms of their judgments of the percentage of useful items retrieved. The IIDA trained participants indicated, on the average, that 46.6% of the references retrieved were useful, while the conventionally trained participants indicated that 62.9% of the references were useful.

A total of 84% of the participants in each of the groups were of the opinion that they would be able to search on their own, and even more were willing to recommend the system to their friends, with 92% of the conventionally trained group, and 88% of the IIDA trained group, checking this question in the affirmative. Neither of these differences are statistically significant. In addition, with thirteen of the IIDA trained, and ten of the conventionally trained, having done some previous research on the topic of the search being rated, there was no difference between the two groups on this measure of prior familiarity with the topic. A final measure on which there was no significant difference between the two groups was the amount of time required to complete the search. At the time the searches were done, the IIDA assistant on site recorded an estimate of how long it had taken to do the search. The mean time to complete for the IIDA trained group was 32.6 minutes, for the conventionally trained group, 30.4 minutes.

The two groups did differ, however, in their evaluation of the search system. The conventionally trained group was significantly more positive, with an average sum across the five scales of 19.6, than was the IIDA trained, with an average sum of 17.8 on the five scales (Mann-Whitney $U = 203.5$, $p < .05$). Inspection of the scores on each of the five scales indicated that the conventionally trained group had given a numerically greater degree of agreement on each scale than had the IIDA trained group. The two groups were compared with each other on each item individually. The conventionally trained gave a significantly greater degree of agreement than the IIDA trained to the statement about the instructions being easy to follow (Mann-Whitney $U = 218$, $p < .05$), and to the statement that the system was easier than expected to use (Mann-Whitney $U = 216.5$, $p < .05$). Further, although the difference was not significant (Mann-Whitney $U = 226$, $p < .07$), the conventionally trained group appeared to give a higher degree of agreement with the statement that the search system was stimulating to use.

The intercorrelations of several post-search measures, the diagnostic categories, and some of the measures from the pre-search questionnaire are reported in Table 7 for the IIDA trained searchers, and in Table 8 for the conventionally trained searchers. Some variables which might have been included in the Tables were omitted for various reasons. For example, the variable of previous research on the topic of the search was left out of the

	1	2	3	4	5	6	7	8	9	10
1. Computer Use	1.00									
2. Degree Year	.17	1.00								
3. Satisfaction	.12	.44*	1.00							
4. Evaluation	.27	.47*	.68**	1.00						
5. Understanding	.43*	.37	.20	.37	1.00					
6. Time	-.06	-.18	.21	-.01	.22	1.00				
7. Global Faults	-.08	-.32	-.16	-.07	.18	.41*	1.00			
8. Local Faults	.10	.06	.05	.12	.52**	.59**	.30	1.00		
9. Syntax Errors	-.33	-.28	-.25	-.30	-.10	.36	.36	.18	1.00	
10. % Useful Items	.13	.33	.75**	.52**	.10	.12	-.46*	.06	-.13	1.00

*p $\leq$ .05
**p $\leq$ .01

Table 7. Correlations for IIDA trained group.

	1	2	3	4	5	6	7	8	9	10
1. Computer Use	1.00									
2. Degree Year	.03	1.00								
3. Satisfaction	.49*	-.04	1.00							
4. Evaluation	.59**	-.15	.60**	1.00						
5. Understanding	.37	.12	.42*	.35	1.00					
6. Time	.00	.10	.34	.17	.39	1.00				
7. Global Faults	.26	.20	.19	.31	.20	.24	1.00			
8. Local Faults	.18	.28	.00	-.19	.08	.06	.09	1.00		
9. Syntax Errors	-.08	.35	-.21	-.20	.01	.08	.26	.18	1.00	
10. % Useful Items	.12	.13	.43*	.33	.29	.26	.12	.04	-.02	1.00

*$p \leq .05$

**$p \leq .01$

Table 8. Correlations for conventionally trained group.

Tables because of a lack of significant correlations with other variables. Similarly, the number of computer languages known, and the self-ratings of computer skill, which were highly correlated for both groups, were omitted because they showed no significant correlations with any of the other variables investigated. Conversely, the judgments of usefulness were omitted because of their high correlations with, and only with, the satisfaction ratings score, and the other variables that satisfaction is correlated with in the Tables. Finally, willingness to recommend the system to others, and the user's belief that he could search on his own were omitted because of the limitation on the value that the point-biserial correlation can assume when most of the responses fall into only one of the two categories on the dichotomous variable.

While Tables 7 and 8 report several correlations that are significantly different from zero, those of major interest are the correlations where either the difference between the correlation in one group and in the other is significant, or where each of the correlations is significantly greater than zero, and there is no significant difference between the two correlations. There are two pairs of correlations which fall into this latter category. In both groups of users, the satisfaction scores correlate significantly and positively with the evaluation ratings, and the judgments of the percentage of useful references retrieved. While the evaluation ratings correlate significantly with the percentage of useful items in the IIDA group, the correlation is not significant in the other group. In addition, the two correlations do not differ significantly from each other. If, however, the two correlations, not being different, are assumed to be two different estimates of the population correlation and are combined, the resulting estimate of the population correlation is significant ($r = .42$, $p < .05$). This pattern of correlations is compatible with some of those reported in the baseline study discussed earlier. There, the possibility was suggested that, for users actually doing their own searching, the degree to which the users are able to retrieve usable results influences their satisfaction with, and evaluation of, the search assistance system. That suggestion would seem to find further support in the patterns of correlations described here, in that the same relationships seem to exist in the data from both studies.

In three of the cases reported in Tables 7 and 8, the correlation for one group differs significantly from that for the other group. The correlations between the year in which the bachelor's degree was earned and the evaluation of the search system differ significantly from one group to the other ($z = 2.26$, $p < .05$), as do the correlations between the time to complete the search and local faults ($z = 2.13$, $p < .05$), and between global faults and the percentage of useful items retrieved ($z = 2.10$, $p < .05$). This latter correlation, which is significant for the IIDA group but not for the conventionally trained group, is probably the result of something specific to the sample of participants used in this study. These same two variables are not correlated significantly in the IIDA trained group described in the baseline study reported above.

In the case of the first two correlations, where the IIDA group differs significantly from the conventionally trained group, both relationships seem to have no immediately obvious explanation, but may represent ideas worth following up in future research. For example, one speculation about the first correlation--between degree year and evaluation for the IIDA but not for the conventionally trained group--is that the IIDA group had considerably more exposure to the total set of IIDA programs, and the more recent graduates, having broader contact with and understanding of the strengths and limitations of computerized systems in general, were more tolerant or forgiving of the limitations of the system.

3.2.4 Responses of IIDA trained users to open-ended questions. On the post-search questionnaire filled out by all participants, there were questions which asked for the user's written comments about various aspects of the search assistance system. This section summarizes the responses of those users who were both trained and assisted by IIDA

a) Positive impressions. When asked to give their positive impressions of IIDA, a large majority of the users directed their comments to various aspects of the assistance mode. Most of these users appreciated the help cues following the error messages. As one user explained it, "I had the feeling that if I goofed up the computer would 'know' how to get me out of it." Some of the other positive impressions about the assistance mode: the help mode allows one to operate more efficiently; help allows one to continue smoothly after making an error; the system did not "over-manage" the search; the error analysis was reasonable, and the "quick

information" notifies the user when he is on a dead-end route, or if the items requested are not in recent indexes.

Several users were impressed with the rapid access to information, and IIDA's method of providing the user with "step wise acquisition of information." Others found IIDA easy to use, and found it easy to understand the command descriptions. A few users liked the way they were guided through the use of the commands, or the search summary option, or the relevance reports about the records retrieved. Overall, IIDA was described as doing "a good job of introducing techniques," with, "helpful instructions," and having assistance available. One user noted that with, "...little practice one could probably become quite proficient." While another reported, "I would use this system often (if available)."

Some participants' comments were directed toward DIALOG or on-line searching in general. Some of these favorable comments dealt with: working with unions and intersections of sets, the ability to change strategy in mid-search, the helpfulness of being able to expand tables, the desirability of being able to retrieve abstracts, and the immediate satisfaction from seeing the results on-line. Some of the mechanical difficulties which plagued the system showed up in the positive impressions in that one user noted that, "DIALOG was useful, when reached efficiently."

b) Negative impressions. When asked to give their negative impressions of IIDA, the users' most frequent responses concerned the limitations or irrelevance of the data base and the numerous telephone communication interruptions. Several users would have preferred a chemistry or physics data base. One user described the "service factor" as miserable because he returned six times for three appointments. Concerning IIDA's instructions, some of the participants found it difficult to assimilate all the information in the exercises, and would have preferred to have a supplementary primer for commands and short cuts. One suggested having shorter sessions, as he found it difficult to work on the system for long intervals. Other users would have liked more time on the system, specifically exercise three, to better understand its use for more efficiency. One user did not think that the system gave enough time to enter commands while the user was trying to remember them.

Some users complained that IIDA's text was too conversational, had too many and too lengthy error messages, and felt that the system took too

long between responses. According to some of the users, the relevance comments were a nuisance; they would have preferred not to give relevance ratings after every abstract. Questioning the validity of the relevance data, one user noted that one's opinion could change of a particular reference as one digs deeper into the subject.

Several individual negative impressions were reported: confusion about which data base to use in exercise two, IIDA not being responsive to minor typing errors, difficulty in aborting a search if it is fruitless, and IIDA could have been more helpful with assistance on problems.

Concerning DIALOG, some users found some items they retrieved were irrelevant to the selection. One user complained that the data base was poorly filed by author listings, while another found 'typos' in the file misleading. Confused by DIALOG's filing system, one user misunderstood the difference between selected phrases and combined single terms.

c) IIDA did the wrong thing. When asked to list the situations in which IIDA did the wrong thing, most comments referred to telecommunications problems: lost data from DIALOG connection requiring additional input, being logged off before it was requested, having the system print a statement five consecutive times. One user noted receiving no response from a command, then receiving the message, "DIALOG no longer responding to commands."

The following are some individually reported situations in which IIDA did the wrong thing. One user began the search with /HELP. After the initial eight frames and instructions to choose a number of a DIALOG command, he entered /HL.52. This is not a DIALOG command, but IIDA returned the user to the main program just as if he had entered a DIALOG command. IIDA asked another user to type in a strange format number (ED/12126/90). An error message appeared each time she tried to enter the number. Toward the end of a search, another user received the "illegal procedure condition" message for no apparent reason. One user mistakenly chose to search in file 201, but could not change files once she had discovered the problem. The system had trouble completing another searcher's references. He abandoned his efforts after several trials. Another user complained that IIDA was not "forgiving" of the spacing between a number and the command (e.g., BEGIN). Yet another user had trouble with the error messages. Three errors, with codes, appeared with instructions to type in a code number to get more information. None of

the code numbers given the user exceeded 250. The user followed instructions, and received a message that acceptable codes included only those between 400 and 500. These mostly represent software errors of omission, i.e., the system was not planned to respond in these ways, but obscure errors caused these peculiarities

d) Search without IIDA. The participants were also asked if they felt they could search on their own, and if so, when they had decided they knew enough to do so. Twenty of the users in this group responded with written comments. Two reported that they could not search without assistance. Four reported deciding they could search on their own after exercise two; one, after exercise three; and one, during the first search. The largest number of users, eight, reportedly decided after the first search was completed. One user decided during the second search; and two, after. Only one person seemed unsure, responding with a "maybe."

When asked to explain their reasons for feeling that they could or could not search on their own, a few users felt they could as a result of gaining confidence from successfully completing a search in the assistance mode with limited assistance. Several of those who responded reported that, while they felt they could search without assistance, they needed a better summary of commands, or that they could search without assistance only on familiar topics. Those who felt they could not search without assistance listed their reasons as not feeling comfortable enough with advanced commands, or needing more details on search strategy, or simply needing more experience. One user did not think of IIDA as being more than the instructional exercises, apparently, since he answered that he could search on his own, "...with help aids."

3.2.5 Responses of conventionally trained users to open-ended questions. This section of the report summarizes the responses of the conventionally trained, but IIDA assisted, users to the questions on the post-search questionnaire calling for written responses.

a) Positive impressions. When asked to give their positive impressions of IIDA, most of the users in the conventionally trained group actually responded to characteristics of DIALOG rather than of IIDA. The largest number of responses about IIDA concerned the help library and the directions given by IIDA following syntax errors. Other responses referred to IIDA's instructions. The comments included: "self-explanatory system,"

"the language is clear," "the system guides the user," "commands were simple," and, "(the system) was enjoyable to interact with."

Overall, the majority of the responses were positive impressions about the time-saving aspect of searching on-line without an intermediary, and the quick responsiveness of the system. Several individual impressions about DIALOG referred to such things as: keeping track of sets accumulated, expand tables offering options to create new sets, having the ability to scan titles quickly, selecting with infixes, combining and sorting sets, and changing commands easily. One user found the system very easy to use, straightforward, and reported that, "It did nothing unexpected."

b) Negative impressions. When asked to give their negative impressions of IIDA, many of the users again responded to characteristics of searching in general rather than to IIDA. Some of the IIDA-specific comments included one user who was dissatisfied with the time limit to either enter a command or be logged off. Another thought the comments made by IIDA were too lengthy and complicated, and a second reported having trouble interpreting them. Three of the users were dissatisfied with IIDA's limitation on the number of records which one is permitted to view with each use of the type command. An additional complaint was made by one user who felt that there was not, and should be, an alert for such simple problems as improper use of the carriage return.

Overall, one user thought the system too complicated, and some of the others thought the whole process of searching took too long, or at least longer than expected. Others referred to the delays and problems with getting on the system, and the pauses that occasionally occurred throughout the search, as a waste of time. Also related to the use of time, one user felt that his time could be used more effectively by doing a preliminary literature search followed by having an exhaustive search performed by a professional searcher. One user would have liked to have had the option of breaking from a print-out of the abstracts after having looked at the first few lines.

In doing the search, some users had trouble selecting key words, since there seemed to be a mismatch between the user's descriptors and those in the file. This problem caused some searchers difficulty in narrowing down the search. Several users were dissatisfied with the selection of available data bases, and a few were unhappy with the data base used. For example,

one reported misspellings filed in the data base, and another was dissatisfied to find Exxon production research not included in the corporate sources.

c) IIDA did the wrong thing. When asked to list the situations in which IIDA did the wrong thing, users reported a wide variety of problems. Some of these problems seem to be related to IIDA specifically, but many of them appear to be based upon telecommunications difficulties and related problems. The kinds of problems users noted included: delays in logging on, being logged off too soon, having to return for numerous sessions because of system communication breakdowns, the system not accepting a command, receiving "erroneous results which neither I nor the IIDA assistant understood," losing the response to a SELECT, and being rejected when trying to re-enter the command, and useless repetitions as the result of "noise." One example of what may have been an on-line noise problem involved selecting a particular descriptor and receiving a null set. When the same descriptor was selected again, it retrieved several items.

d) Search without IIDA. The participants were also asked if they felt they could search on their own, and if so, when they had decided they knew enough to do so. Twenty-one of the conventionally trained users felt they could search without assistance. Three reported first deciding this after the conventional training course, and one had already decided before the training began. Two of the users decided during their first assisted search, and seven, after. Another six decided during the second search, and two reported deciding after the second search. The major reason for feeling they could search without assistance was the simplicity of the commands. Some thought they would still use the course training manual. One thought he could search without any further assistance or reference, but not in the engineering file or in fields outside of his interests.

The four users who felt that they would be unable to search without assistance reported that they needed help defining the search terms, or in finding appropriate terms. One user thought that he simply needed more experience.

3.3 Discussion

In this study, two randomly assigned groups of users provided feedback on the results of an on-line bibliographic search. Both groups of users did the search themselves with the assistance of the IIDA on-line

assistance programs. One group of users had been trained in searching by the IIDA instructional exercises, while the other group of users received their training by a more conventional training method. Comparison of the two groups on several pre-search measures revealed only one significant difference between the two groups. This difference, in the average number of computer languages known by the users in the the two groups, did not seem to be a very strong one, however, in that further investigation indicated that the number of users who knew no languages at all was not significantly different between the two groups, and that those users who knew one or more computer languages did not differ between the two groups in terms of the mean number of languages known. Further, the number of languages known turned out not to be correlated with any other variable investigated in the study except the individual's self-rating of computer skills.

The two groups also did not differ from each other on several post-search measures of their reactions to the searches and the search assistance system. For example, in terms of judgments of the percentage of useful items retrieved in the searches, the difference was not statistically significant. Where differences did exist, they were related to characteristics of the search assistance system rather than the search results. Those users who were conventionally trained seemed to be more likely than the IIDA trained searchers to agree that the instructions in the search system were easy to follow, and that the system was easier to use than had been expected. Further, the conventionally trained group expected to encounter more difficulty with the system than was in fact the case. These differences reflect favorably on the IIDA assistance programs in that experience with the IIDA training exercises is clearly not necessary for the individual to be able to make use of the assistance mode.

It was also found that there were no significant differences between the two sets of searches in terms of the number of IIDA detected errors or faults in the searching behavior of the two groups. Finally, although statistical analyses were not conducted, the responses of the two groups of searchers to open-ended evaluation questions seemed to clearly indicate two things. The first is that IIDA is not a finished product without fault. The questions designed to elicit negative comments from the users were clearly able to do so. The second is that, overall, most of the users

seemed to like learning to search, either with or without IIDA, and they seemed quite positive about the IIDA assistance they received. It is also interesting to observe that the users did not draw a clear distinction between the characteristics of IIDA and those of DIALOG or the telecommunications. This is not surprising, since IIDA was designed to be as unobtrusive as possible during searching. However, system designers should be aware of the fact that users generally tend to evaluate the system as a whole, and that the newly designed component, such as IIDA, may get either the credit or the blame for characteristics of the other parts of the system.

In light of the intent of this study, i.e., to compare IIDA training with more conventional training, with a view to refinement of the IIDA programs, it was surprising to discover so few significant differences between the two groups of searchers. Further, the two groups, trained by different methods, did not differ appreciably either on the diagnostic measures or in the percentage of useful references retrieved. A significant difference on one or more of the diagnostics would have pointed toward either a deficiency in the IIDA training materials, or in the usefulness of the diagnostic messages during assisted searching. This would have led to a revision of one or both aspects of the system. The original design specification for IIDA did not envision a system to be competitive with more conventional training approaches. Rather, the intent was to provide an avenue of access to on-line searching for individuals who could not or would not do a conventional training course, and who still wanted personal rather than intermediated access to a data base.

The finding of no significant differences on the diagnostic measures could pose a problem for interpretation of the results were it not for the fact that both groups were able to retrieve a significant percentage of useful references during their searches. Presumably this represents something which neither group would have been able to accomplish if simply turned loose with a terminal, but without having any training or assistance of any kind. Further, the two groups did not differ in their estimates of the percentage of useful references retrieved. This pattern of results argues strongly for the idea that IIDA training, as it is presently structured, represents a viable alternative to the type of conventional training with which it was compared. One reasonable guess as to the reason

for the pattern of results reported here is that, while the human instructor may well in fact have been more flexible and responsive in assisting the student and in answering questions during training, the design of IIDA does, as intended, enable the user to discover where and how to go get the information needed to answer questions for himself.

4. General Discussion

A major goal of this project has been to provide a method for allowing direct user access to bibliographic searching. Thus, the attempt has been to develop a set of computer software packages which can provide on-line assistance to occasional users of retrieval systems. This collection of programs is also intended to be able to provide instruction, if needed, in the commands used in searching and in search strategy. When originally conceived, the expected utility of IIDA lay in the area of what might be referred to as "problem solving searches." These are searches where the end user of the information does not know exactly what the characteristics are of the desired set of references until they have actually been found. Consequently, it is very difficult for the end user to describe the problem to an intermediary. There is no reason, however, why the IIDA user could not and should not make use of the system for all kinds of searches, if desired.

The primary intended IIDA user has been the working scientist or engineer who may need access to the data base only a few times a year, and consequently is not interested in training oriented toward those who become professional intermediaries. This person is assumed to be comfortable using computers, but not necessarily trained in their use. In addition, it is assumed that this user is a serious, well-intentioned searcher who is trying to use the system to solve a problem. The IIDA software and diagnostic procedures were created to help remove the barriers to access for these users. Minor modifications of the system, however, could make it available to a much wider audience.

Given the nature of IIDA, it is not possible to completely separate the performance of IIDA from the performance of the IIDA users. Looking first at the baseline study, the data suggest that IIDA users did searches which were, on two kinds of measures, done as well as those done by the more conventional approach of having the searches done for them by a professional searcher. First, in the judgments of the percentage of useful references retrieved in the searches, there was no significant difference between the results obtained by the professional searcher and those obtained by the user performing his own search. Second, it was found that there were no significant differences in the number of IIDA detected errors or faults between the newly trained, IIDA assisted searchers and the professional

searchers. Together, these findings imply that the results of the searching were about equal, and that the performance of the searchers was about equal. In addition, the average rated satisfaction of the IIDA users and the participants who had intermediated searching done did not differ, and the majority of the users in both groups were pleased enough with their results and experience that they were willing to recommend the system to their friends.

The two different user groups which differed in the method of training prior to doing IIDA assisted searching also did not differ appreciably either in the diagnostic measures or in the percentage of useful references retrieved. In addition, the members of both of these groups did not differ significantly from each other on several measures of their reactions to the search system, with, for example, most of the participants feeling they would recommend the system to their friends. Where the two groups did differ, however, was in the evaluation of the system instructions and the ease of use of the system. In both cases, the conventionally trained users indicated a higher degree of agreement than did the IIDA trained users. That is, they were more likely to feel that the instructions were easy to follow, and more likely to indicate that the system was easier to use than they had expected.

The finding of so few significant differences, particularly on the diagnostic measures, came as somewhat of a surprise, in that IIDA was designed to produce acceptable results for its users, but it was never expected that the step-by-step performance of the users would match that of professionals. Also, the original design specifications of IIDA did not envision a system which would be competitive with more conventional training approaches. Rather, the intent was to provide an avenue of access for individuals who could not or would not do a conventional training course, and who still wanted personal rather than intermediated access to a data base.

In a case where no significant differences exist, the situation is always troublesome in that such a finding could occur because the two groups did not differ, or because the measures chosen were irrelevant. This latter explanation, however, seems to be less reasonable than the idea that the system is working as it was designed to work. One reason for arguing that the training and diagnostic assistance routines worked as intended lies in the fact that in the baseline study the end user evaluation of the utility

of the information retrieved in the searches did not differ significantly between the two groups. This means that a group of individuals who had never before done on-line searching were able, with IIDA training and assistance, to do searches which produced results containing as much useful information as was contained in the results of searches done by professional searchers. A second reason for not assuming that the diagnostics are irrelevant as indices of the searching behavior of the various user groups lies in the fact that the diagnostics were all empirically developed. That is, they were all designed to index and deal with problems encountered by searchers which have been either observed by the designers, reported in the literature by others, or both.

However, recognizing that the conclusions drawn about the effectiveness of the diagnostics are not as clear and unambiguous as might be desirable, there are some additional directions for future research which ought to be followed. The first of these has to do with the evaluation of the diagnostics. The overall requirements for evaluation of the project as a whole, as well as the available resources, dictated that some studies have a higher priority than others. This led to a concentration on the evaluation of the IIDA system as a system rather than a more direct focus on the diagnostic component of the system. A future evaluation devoted solely to the question of the performance of the diagnostic system, with and without various kinds of training in advance, would be both useful and desirable.

Another line of investigation dealing with the diagnostics which could be followed has to do with the possibility of adaptation to user skill level. The diagnostics were designed for use in a limited context: the training and assistance of the kind of novice user described earlier and in the user studies reported here. There was some indication in the experiences of the professional searchers that, when the diagnostics are employed with highly experienced searchers, they may not be the most appropriate procedures. There may have been times when the diagnostics were triggered by relatively sophisticated searcher behavior which is appropriate in one context, but not in situations with which IIDA was designed to cope. Thus, there seems to be no reason to assume that the same diagnostics will be equally useful or desirable for all user skill levels. Consequently, the diagnostics should be tested out more thoroughly

with persons of differing skill levels. In addition, it would also be desirable to test different variations of the diagnostics. In particular, it would be of interest to determine whether or not it is possible, by varying the thresholds in the existing set of diagnostics as a function of the prior performance of the user, to get the diagnostics to perform adequately with individuals of different skill levels. That is, can they be made adaptive?

A third direction for future research has to do with the type of IIDA user. Although the original target user group for IIDA consisted of technically trained individuals interested in a particular class of search problems, there now seems to be no reason not to attempt to extend IIDA. The technically or scientifically trained user may be only one of several kinds of users who would find IIDA attractive and useful. In particular, it seems desirable to determine whether or not a system such as IIDA can be used to provide direct data base access for a wide variety of possible end users interested in a wide variety of search problems. In fact, it may well be that IIDA, or an IIDA-like system, is a viable method of introducing the computer and on-line searching to users with little or no background with computers.

The fourth recommendation for future research follows from one of the limitations of the studies described above. These studies, and many other studies on searcher behavior, tend to be flawed by the fact that there is a relatively short amount of time between when the participants in the testing are first exposed to the system, and the evaluation of their performance, or the system performance. In addition, there seems to have been little or no testing of information system users at several points over a long period of time in the investigation of development or change in behavior over time. In the studies described above, the volunteer subjects generally tended to want to do the whole set of exercises, and then do their searching in a relatively short time. It seems particularly desirable that a long-range study be done, over a period of a year or so, in which the focus is upon how different people adapt to a new system, how quickly they adapt, how their behavior changes over time, and how it changes as the result of multiple search experiences.

In conclusion, it seems that a new idea has been fairly tested in the very environment for which the concept was intended, and has come off quite

well. Indeed, one of the important characteristics of the two studies reported here is that there appeared to be no differences among the results produced by the various user groups when there was every reason to expect in advance that there would be a number of differences. While some aspects of these results are not entirely conclusive, they are supportive of the idea that the IIDA diagnostic procedures did indeed adequately measure important aspects of user performance. What is more certain, however, is that the IIDA system represents a way of training and assisting novice users in doing their own data base searching on problems of their own choice, with a level of performance that matches that of more experienced searchers. Furthermore, IIDA clearly represents a viable alternative route to gaining direct data base access for those end users who can not or will not do more conventional forms of search training.

5. References

1. Individualized Instruction for Data Access, Quarterly Report No. 6, NSF Grant No. DSI 77-26524, School of Library and Information Science, Drexel University, Philadelphia, Pa., September, 1979.

2. Individualized Instruction for Data Access, Quarterly Report No. 7, NSF Grant No. DSI 77-26524, School of Library and Information Science, Drexel University, Philadelphia, Pa., December, 1979.

3. A Guide to DIALOG Searching, Lockheed DIALOG Information Retrieval Service, Palo Alto, Ca., 1979.

4. Individualized Instruction for Data Access, Quarterly Report No. 8, NSF Grant No. DSI 77-26524, School of Library and Information Science, Drexel University, Philadelphia, Pa., March, 1979.

6. Appendix

APPENDIX A: Questionnaires

Bibliographic Searching

Pre-Search Questionnaire

Name: _____

Code Word: _____

(Note: Your responses to this questionnaire will be confidential and will not be made public except in a group or aggregate form which will not allow for the identification of any individual's responses.)

Individualized Instruction for Data Access (IIDA) Project

Drexel University

Philadelphia, Pa.

We would like to have you give us some information about yourself so that we have a better idea of the needs and interests of our users. On the following pages you will find several questions. Some call for a written answer, others, for a check mark indicating your choice among several alternatives. Even if you find some of the questions strange or inappropriate it is important that you complete them all.

1. What is (was) your undergraduate major in college? (Be as specific as possible) _____

2. Year undergraduate degree received or expected. _____

3. Do you have any graduate degrees? Yes: _____ No: _____

   If yes, what degree(s) in what area(s)? (e.g., M.S. in Psychology)

   _____

   _____

4. Please give a brief description, with field and title, of the full-time jobs you have held during the last five years.

   _____

   _____

   _____

   _____

5. Please give a brief description, with field and title, of the full-time job you would most like to be working at five years from now.

   _____

   _____

   _____

   _____

6. Do you know any computer languages? Yes: _____ No: _____

   If yes, which one(s)? _____

7. How would you describe your computer programming skills?

_____ :Non-existant
_____ :Poor
_____ :Good enough to get by
_____ :Quite good
_____ :Excellent

8. Have you ever had anyone do on-line computerized bibliographic searching for you?

Yes:_____ No:_____

If yes, roughly how many times or how often? _____

9. Have you, yourself, ever done any on-line searching?

None:_____ Some:_____ A lot:_____

When you have identified a work-related problem, how do you identify the information you need to solve the problem? That is, to what extent do you rely on the following individuals or resources?

Asking co-workers:

:_____:_____:_____:_____:_____:
not at all very little sometimes a moderate amount considerably

Asking professional colleagues:

:_____:_____:_____:_____:_____:
not at all very little sometimes a moderate amount considerably

Asking your supervisor:

:_____:_____:_____:_____:_____:
not at all very little sometimes a moderate amount considerably

Going to the library:

:_____:_____:_____:_____:_____:
not at all very little sometimes a moderate amount considerably

Asking the information retrieval specialist:

:_____:_____:_____:_____:_____:
not at all very little sometimes a moderate amount considerably

Solving a research or design problem often requires several steps, some of which may be repeated several times before the problem is solved.

These typically involve:

a) recognizing the problem exists
b) defining the problem
c) breaking the problem up into sub-problems
d) selecting one of the sub-problems for "solution"
e) generating options
f) selecting an option
g) implementing the option
h) evaluation of the result

At what point(s) in the problem-solving process is(are) bibliographic materials most useful?

At what point(s) have you typically used bibliographic searching?

At what point(s) can you imagine yourself making use of these resources?

At what point(s) would you recommend to others that they try bibliographic searching?

Intuition should play very little role in solving problems:

:_____:____:_____:_____:_____:
Strongly agree Agree Uncertain Disagree Strongly disagree

I do not like using the computer for work-related activities:

:_____:____:_____:_____:_____:
Strongly agree Agree Uncertain Disagree Strongly disagree

I believe that every problem should have a clear solution:

:_____:____:_____:_____:_____:
Strongly agree Agree Uncertain Disagree Strongly disagree

One of the benefits of an education is that it helps you to learn how to solve problems:

:_____:____:_____:_____:_____:
Strongly agree Agree Uncertain Disagree Strongly disagree

There should be a set or routine procedures to follow in solving problems:

:_____:____:_____:_____:_____:
Strongly agree Agree Uncertain Disagree Strongly disagree

I believe that using a computer for work-related projects will be:

complex:___:____:_____:_____:_____:____:___:simple
very quite slightly neither slightly quite very

boring:___:____:_____:_____:_____:____:___:interesting
very quite slightly neither slightly quite very

painful:___:____:_____:_____:_____:____:___:pleasurable
very quite slightly neither slightly quite very

unsuccessful:___:____:_____:_____:_____:____:___:successful
very quite slightly neither slightly quite very

worthless:___:____:_____:_____:_____:____:___:valuable
very quite slightly neither slightly quite very

frustrating:___:____:_____:_____:_____:____:___:not frustrating
very quite slightly neither slightly quite very

Bibliographic Searching

Intermediate Questionnaire

Name: _____

Code Word: _____

(Note: Your responses to this questionnaire will be confidential and will not be made public except in a group or aggregate form which will not allow for the identification of any individual's responses.)

Individualized Instruction for Data Access (IIDA) Project

Drexel University

Philadelphia, Pa.

How do you feel about the amount of effort you made during the search?

very dissatisfied	dissatisfied	satisfied	very satisfied

How do you feel about the amount of time you spent in the process?

very dissatisfied	dissatisfied	satisfied	very satisfied

How do you feel about the kind of assistance you received from the system?

very dissatisfied	dissatisfied	satisfied	very satisfied

How do you feel about your total search experience?

very dissatisfied	dissatisfied	satisfied	very satisfied

How do you feel about the items you retrieved?

very dissatisfied	dissatisfied	satisfied	very satisfied

How do you feel about the search as a whole?

totally useless	useless	useful	extremely useful

How do you feel about the results of the search itself?

not at all useful	moderately useful	very useful--I got what I needed

Have you ever done any previous research on the topic of your search?

Yes: _____ No: _____

What percentage of the items retrieved would you say were:

Very useful: _____
Useful: _____
Useless: _____

Total: 100%

Would you recommend this system to your friends:

Yes: _____ No: _____

Please describe your overall impressions of working with computer assisted bibliographic searching.

Positive Impressions: (What did the system do or say that you liked most?)

Negative Impressions: (What did the system do or say that you liked least?)

Please describe any situations where the system did the wrong thing.

Bibliographic Searching

Post-Search Questionnaire

Name: _____

Code Word: _____

(Note: Your responses to this questionnaire will be confidential and will not be made public except in a group or aggregate form which will not allow for the identification of any individual's responses.)

Individualized Instruction for Data Access (IIDA) Project

Drexel University

Philadelphia, Pa.

How do you feel about the amount of effort you made during the search?

very dissatisfied	dissatisfied	satisfied	very satisfied

How do you feel about the amount of time you spent in the process?

very dissatisfied	dissatisfied	satisfied	very satisfied

How do you feel about the kind of assistance you received from the system?

very dissatisfied	dissatisfied	satisfied	very satisfied

How do you feel about your total search experience?

very dissatisfied	dissatisfied	satisfied	very satisfied

How do you feel about the items you retrieved?

very dissatisfied	dissatisfied	satisfied	very satisfied

How do you feel about the search as a whole?

totally useless	useless	useful	extremely useful

How do you feel about the results of the search itself?

not at all useful	moderately useful	very useful—I got what I needed

Have you ever done any previous research on the topic of your search?

Yes: _____ No: _____

What percentage of the items retrieved would you say were:

Very useful: _____
Useful: _____
Useless: _____

Total: 100%

Would you recommend this system to your friends:

Yes: _____ No: _____

Please describe your overall impressions of working with computer assisted bibliographic searching.

Positive Impressions: (What did the system do or say that you liked most?)

Negative Impressions: (What did the system do or say that you liked least?)

Please describe any situations where the system did the wrong thing.

The process of searching was similar to what I had expected:

strongly agree	agree	uncertain	disagree	strongly disagree

The search system was characterized by instructions that were easy to follow:

strongly agree	agree	uncertain	disagree	strongly disagree

The search system was frustrating to use:

strongly agree	agree	uncertain	disagree	strongly disagree

The search system was stimulating to use:

strongly agree	agree	uncertain	disagree	strongly disagree

The use of the search system was easier than expected:

strongly agree	agree	uncertain	disagree	strongly disagree

How much understanding of the information retrieval system do you feel you gained through your experience?

none	very little	moderate	considerable	complete

Do you feel you have learned enough about searching that you could search on your own?

Yes:_____ No:_____

If yes, when did you first decide that you knew enough about the process to do it on your own?

Please explain your reasons for feeling that you could or could not search on your own without assistance.